Hourly Data Flow

Dave Stowers, Jet Propulsion Laboratory, California Institute of Technology

Heinz Habrich, Federal Agency for Cartography and Geodesy

1 Concatenation of Hourly RINEX Files

Since the beginning of the IGS in 1992 daily RINEX observation files are used to generate most of the final IGS products (IGS satellite orbits, Earth rotation parameters, ...). Using the daily files the observations are too late at the analysis centers available for the generation of "near real time" products, e.g., "ultra rapid orbits" and troposphere parameters. Therefore, currently a subset of the IGS stations submits hourly RINEX observation files additionally to the daily files. Each observation of such sites will be transmitted twice within the IGS data flow. If it is possible to concatenate successfully the hourly files to a daily one, the daily file transfer may be cancelled. Such a procedure would benefit from the more stable transfer of small data files and from the availability of the daily files immediately after the completeness of the 24 files of a day at the data centers. However, it has to be demonstrated that the "original" daily files and the concatenated hourly files include exactly the same observations and site information.

Requirements of Hourly Files

The daily RINEX files may be replaced by the concatenated hourly files if those fulfill the following requirements:

- All 24 hourly files have to be available at the data center.
- The information of the file header have to be identical to that of the original daily file.
- The observations have to be logged continuously and must not produce jumps at the one hour boundaries.
- The phase observations of each hourly file must not be reduced to small numerical numbers (as some RINEX conversion programs did in the past).
- All digits of the numerical data fields have to be identical to that of the original daily file in order to guarantee the same results from the analysis of the two different file types.

A comparison between the daily and the concatenated hourly files for a limited period of time may show the achievement of all the requirements.

Comparison between Daily and Concatenated Hourly Files

In order to remove computer specific differences between the two files the full ASCII RINEX files have to be reconstructed from the "Compact RINEX" format and from the compressed files (assuming *D.Z files in the data base). A character by character comparison as performed by some commands of the computer operating systems (e.g., the Unix command "diff") may not be used, because some differences are unavoidable , e.g., different file creation dates or acceptable, e.g., a blank string is used instead of a leading zero. We have to read each RINEX file following the format definition and may then compare the content of each data field. BKG uses the program "RNXDIFF" for such a comparison.

The Program RNXDIFF

The program RNXDIFF makes use of subroutines of the Bernese GPS Software to get the content of each data field and successively compares the fields of the daily files with that of the concatenated hourly files. For each field a so called "error code" is defined which is set to "1" in the case of an acceptable difference and to "9" in the case of an unacceptable difference. The error codes remain "0" if no differences are detected. The sum of the error codes of all data fields is called "quality code" of the file and is given in the daily summary files of RNXDIFF and in a plot file as station specific time series, too.

Table 1 shows the RNXDIFF summary file of the station WTZR for the day of year 155, 2000. The head of this file includes a list of the data fields that are compared and the definition of the corresponding error codes 1 or 9. In the comparison shown in Table 1 there were differences in two types of data fields detected, namely the "signal noise ratio" and the "epoch" fields. Different signal noise ratios were detected in 11 observations. Because this type of data field has the error code 1 the quality code for the file has been increased by 1. The last epoch was missing in one of the files compared. One missing epoch may be accepted and we defined the error code as 1 for this case. However, if more that one different epochs would be detected the corresponding error code will be set to 9. In the summary file of Table 1 the quality code was increased by 1 because of the different epoch that was detected. After taking into account all types of data fields the quality code results to 2 as given in the last line of the summary file and also indicated in the file name WTZR155**2**.00S. Following the definitions of the quality code we may conclude that the concatenated hourly files may replace the daily files, if the quality code is smaller than 9.

Figure 1 shows all quality codes of the station WTZR as computed from RNXDIFF for the period April, 12 to June 5, 2000. 94 % of the quality codes are smaller than 9 and allow the interpretation that the hourly files of the station WTZR may be concatenated with such reliability.

```
RNXDIFF  V1.0.1 UX                                        04-JUN- 0 19:48

 COMPARSION OF 2 RINEX OBSERVATION FILES
 ---------------------------------------

 IGS DATA CENTER
 BKG, FRANKFURT

 RINEX FILE 1 : WTZR1550.00O
 RINEX FILE 2 : WTZR155C.00O

 CHECKED FOR:

    RINEX FORMAT VERSION              |9| RINEXVER  |
    SATELLITE SYSTEM TYPE             |9| SATSYS    |
    SITE NAME                         |9| SITENAME  |
    SITE NUMBER                       |9| STANUMBER |
    OBSERVER                          |1| OBSERVER  |
    AGENCY                            |1| AGENCY    |
    RECEIVER NUMBER                   |9| RECUNIT   |
    RECEIVER TYPE                     |9| RECTYPE   |
    RECEIVER FIRMWARE                 |9| RECVERS   |
    ANTENNA NUMBER                    |9| ANTNUMBER |
    ANTENNA TYPE                      |9| ANTTYPE   |
    ANTENNA ECCENTRICITY              |9| ANTECCENT |
    NUMBER OF OBSERVATION TYPES       |9| NUMOBSTYP |
    OBSERVATION EPOCHS                |9| EPOCH     |
    NUMBER OF SATELLITES IN EPOCHS    |9| NUMSATEPO |
    OBSERVATIONS                      |9| OBSEPO    |
    SIGNAL NOISE RATIO                |1| SIGNAL    |
    LOSS OF LOCK INDICATOR            |1| LLI       |

 SIGNAL      : ( 0:47:30.000) SIGNAL NOISE RATIO
 SIGNAL      : ( 1:40:  .000) SIGNAL NOISE RATIO
 SIGNAL      : ( 5:24:  .000) SIGNAL NOISE RATIO
 SIGNAL      : ( 5:46:  .000) SIGNAL NOISE RATIO
 SIGNAL      : (12:31:30.000) SIGNAL NOISE RATIO
 SIGNAL      : (13:45:  .000) SIGNAL NOISE RATIO
 SIGNAL      : (17:44:  .000) SIGNAL NOISE RATIO
 SIGNAL      : (22: 4:30.000) SIGNAL NOISE RATIO
 SIGNAL      : (22:57:30.000) SIGNAL NOISE RATIO
 SIGNAL      : (23:37:30.000) SIGNAL NOISE RATIO
 SIGNAL      : (23:50:  .000) SIGNAL NOISE RATIO
 EPOCH       : (23:59:30.000) DIFFERENT EPOCH

 SUMMARY:
 -------
 EPOCHS FILE 1:  2880
 EPOCHS FILE 2:  2879

 COMPARISON QUALITY CODE:
 (0=NO DIFFERENCE, 1-8=NEGLIGIBLE DIFFERENCE, 9=FATAL DIFFERENCE)
 2
```

Table 1: RNXDIFF Summary File WTZR1552.00S

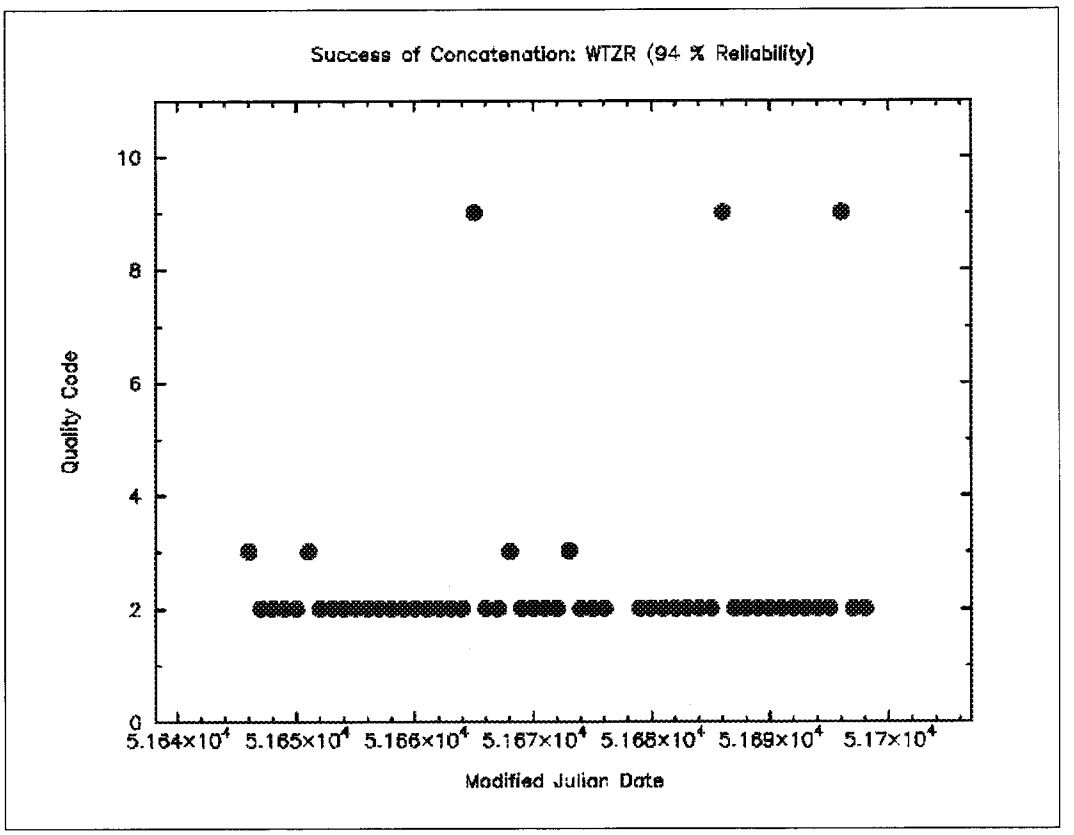

Figure 1: RNXDIFF Quality Code for Station WTZR

Figure 2 shows the reliability as given in Figure 1 for WTZR for all stations that submit hourly data files to BKG. Only 4 stations (KIR0, MAR6,VIL0, andVIS0) reach 100 % reliability, but it has to be taken into account that these stations have submitted hourly files since 3 days. A detailed study of the data flow of each station has to be performed in order to improve the reliability of the concatenation. The transfer of the daily files may be cancelled on condition that the reliability of the concatenation of the hourly files amount to 100 %.

The occurrence of differences separated for each type of the data fields is given in Figure 3. It shows the number of files which were detected to have a difference in the specified data field as the percentage of the total number of files for the period April, 12 to June 5, 2000. In 77 % of the files a difference in the observation epoch was detected. It has to be mentioned that most of the messages were caused by a detected difference in exactly one epoch of the file, e.g., the last epoch in the hourly file of the station WTZR is "23:59:00", but it is "23:59:30" in the daily file. The next most frequent differences are the

"loss of lock indicator (LLI)" and the "signal to noise ratio" data fields. Table 2 shows an example for such type of difference. The GPS satellites SVN numbers 29 and 30 came up as new satellites in the epoch given in Table 2. Whereas the original daily RINEX files shows an LLI and signal to noise ratio of 1, these data fields are left blank in the original concatenated file. The difference seems to be caused by the RINEX concatenation program and will probably not affect the analysis results. It has to be discussed whether such differences may be accepted. In this case the percentage of files showing differences in Figure 3 would be much smaller.

Next Steps

Respond to user input:

"I have been using hourly data for some time and they do have quality control problem. there are corrupted files, files that have duplicated headers, cddis server down, and individual site delay in posting data."

"At the moment we have only some small remarks: data archives:
1. With the growing number of hourly sites a huge number of entries will be in the daily directories. We would propose to insert hourly sub-directories. This has the advantage that you will have not have long lists from ftp if you want to get only the last 1 or 2 hours of data. So our procedures will speed up (every minute counts). At the moment only BKG has hourly subdirectories. What we would like to see is at least a unification between all data centers.

2. It is a problem to mirror the hourly data between e.g. CDDIS and IGN without loosing valuable time. In case of a failure in one data center no immediate switch to another center is possible without loosing data. One solution may be that each site send its data to both data centers any time. Important at least is that all sites automatically switch to another data center if they cannot put correctly their data to the chosen data center.

3. flagging of bad sites
RINEX: There could be one log-file were bad RINEX files are listed. The list could contain informations of the problems (some kind of descriptors), e.g. gap in the data (from ionospheric problems), too few satellites, too many cycle slips, too much multipath, too few obs in low elevations, etc. .These information can in principle extracted from the quality files.

SINEX: A list of problematic sites could be extracted.

All this information are principle available already now. However not all customers and sites will have scripts which monitor all informations. An extract in one table (updated daily, weekly), one table per year /per month or per week will help to get a better overview."

RINEX provider (defined here as the responsible party who determines the useability of the RINEX data) should have final "say" in whether a data center further "upstream" should concatenate hourly files for a particular site on a particular day.

Some sites are providing hourly data in a such a way that the hourly RINEX is not exactly equivalent to what would be represented in a daily file...1s data normal pointed to 30s using an hour of data will likely result in slightly different results, especially on the period's boundary. Re–normal–pointing the 1s data to form a daily 30s product will likely result in fewer discontinuities at the hour boundary that the concatenated file. Pure decimation of the receiver data would solve a portion of this problem, but with a possible increase in data noise.

Consistent hourly data availability from sites contributing hourly data is also problematic. Sequencing of data, delays in transmission, partial hourly files (which may be desirable in the daily file but not in an hourly file) all contribute to problems that make the question of whether hourly data concatenation at the IGS global data center level, highly questionable.

Perhaps if the lower level data centers could transmit a dummy file indicating that a local comparison (hourly concatenation to daily file) has been made and was successful, a higher level concatenation could be performed...but what happens when the GDC performs an adequacy test and determines that there are files which are missing or corrupt? a message back saying please resend certain files/hours? If the OK message from the local data center included the filenames/sizes/crc? of the transmitted hourly data files, this task might not be so onerous. Inter data center equalization mechanisms would assist in this process.

–

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Adminstration.

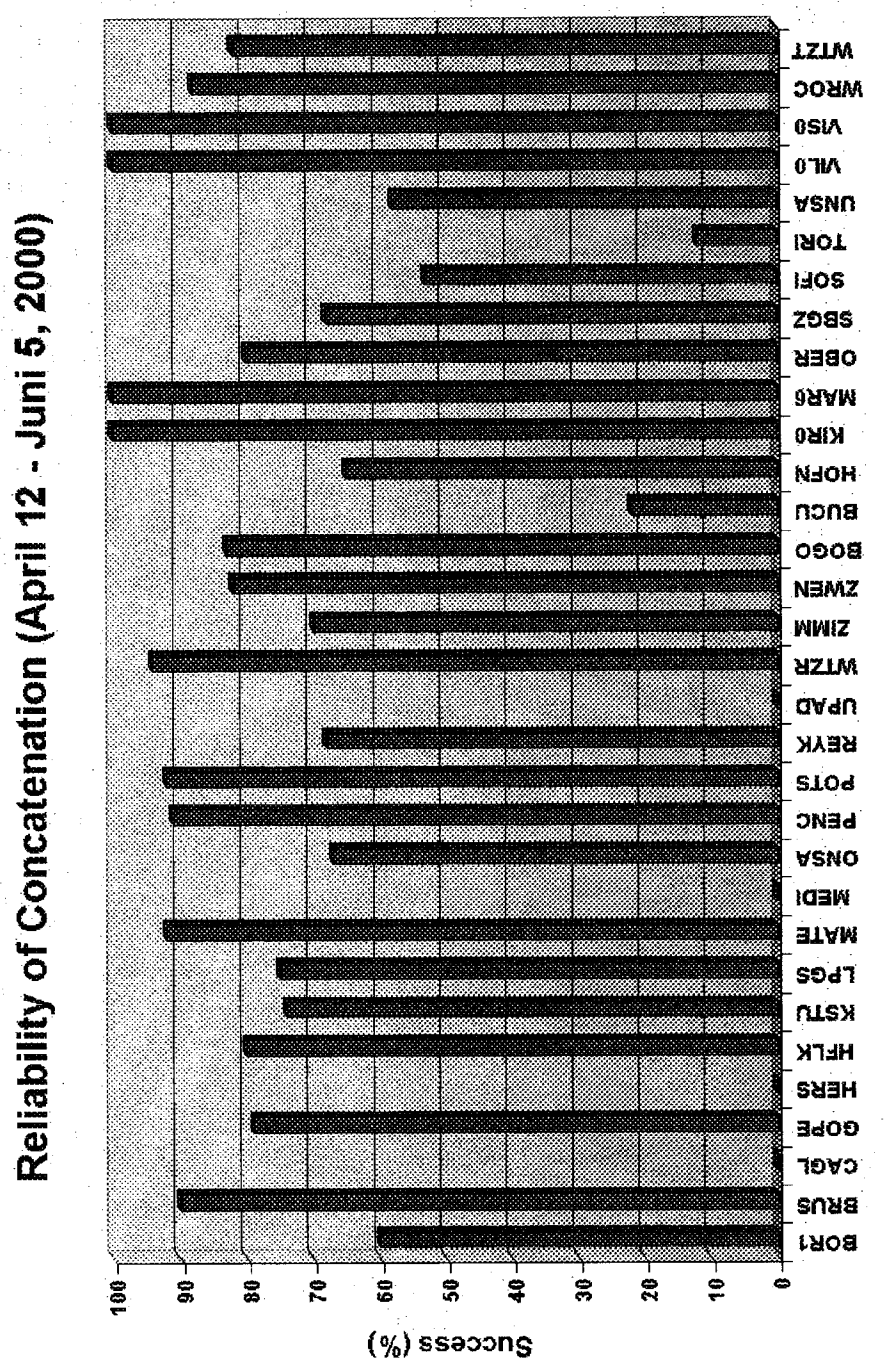

Figure 2: Reliability of Concatenation of all Stations

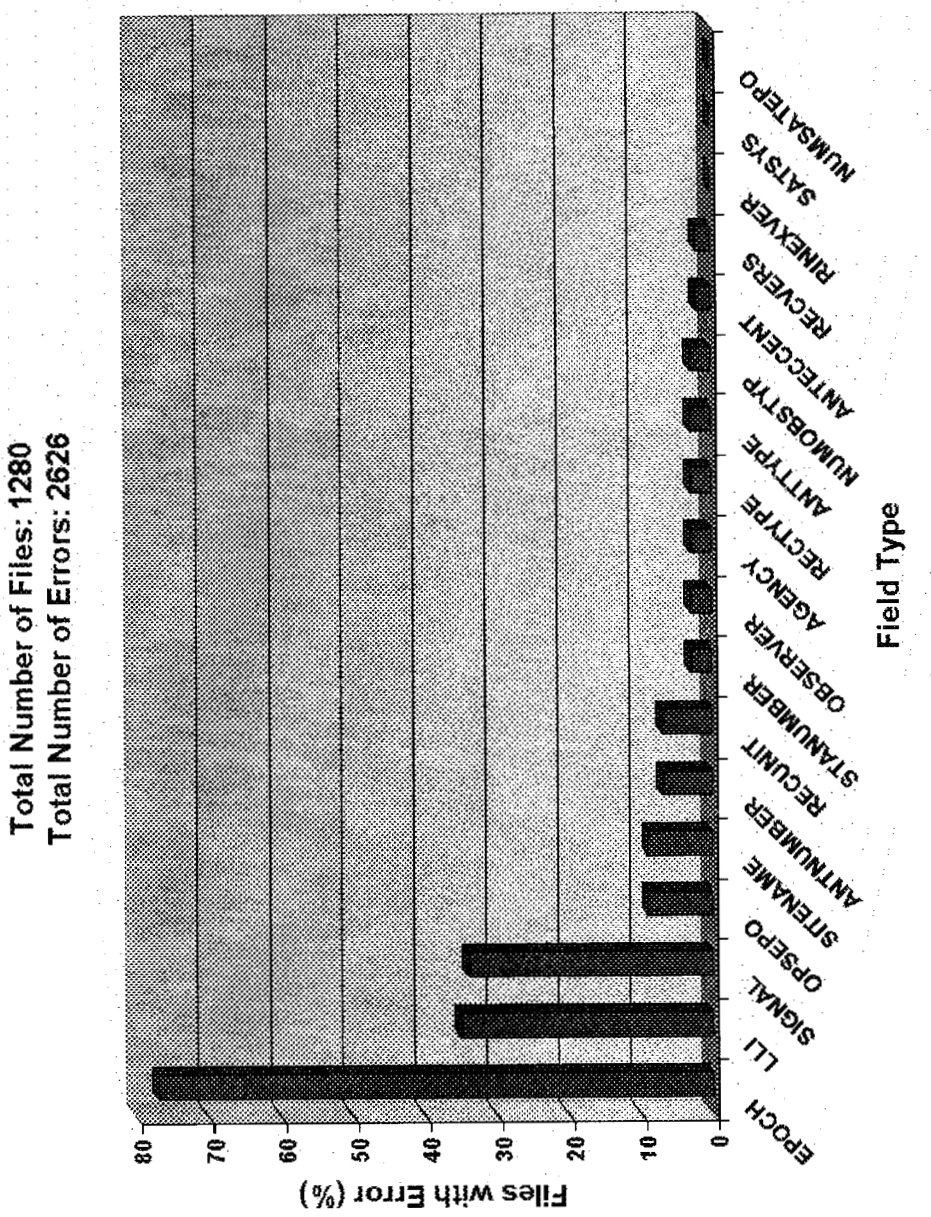

Figure 3: Error Frequency from RNXDIFF

Original Daily File

```
     2              OBSERVATION DATA    G (GPS)             RINEX VERSION / TYPE
GPS-DATA Ver 1.08                                           PGM / RUN BY / DATE
...
00  6  5  0 56 30.0000000  0  9G02G05G07G09G21G23G26G29G30
  24491272.72805  24491277.60405   3793908.16505   3076292.75805      -3255.001
     -2536.366
  22114544.74508  22114550.23108 -14814452.13708 -11359188.94708       2318.174
      1806.366
  22380138.83008  22380144.08108 -10412707.03808  -7965747.02408       -858.687
      -669.107
  20408297.14309  20408301.43309 -20430454.88709 -14777553.80709        214.022
       166.771
  24189931.96305  24189938.55105  -5718742.24405  -4321438.60005        303.064
       236.145
  23466415.78105  23466422.12306  -7157669.21606  -5420564.84006      -1972.672
     -1537.149
  21990865.10009  21990870.04709  -5776160.90809  -4352051.79909      -3197.496
     -2491.555
  24914892.41505          .00001    -51341.91501          .00011       3019.125
           .000
  25477673.83404          .00001    -54512.57701          .00011       3206.270
           .000
```

Concatenated Hourly Files

```
     2              OBSERVATION DATA    G (GPS)             RINEX VERSION / TYPE
CCRINEXO V2.2.2 UX  BKG, Frankfurt/M.   06-JUN- 0 03:25     PGM / RUN BY / DATE
Concatenate RINEX hourly files                              COMMENT
CCRINEXO V2.3.0 UX  LPT                 05-JUN-00 03:04     COMMENT
                                                            COMMENT
GPS-DATA VER 1.08                                           COMMENT
...
-2482.766
  0  6  5  0 56 30.0000000  0  9 02 05 07 09 21 23 26 29 30
  24491272.728 5  24491277.604 5   3793908.165 5   3076292.758 5      -3255.001
     -2536.366
  22114544.745 8  22114550.231 8 -14814452.137 8 -11359188.947 8       2318.174
      1806.366
  22380138.830 8  22380144.081 8 -10412707.038 8  -7965747.024 8       -858.687
      -669.107
  20408297.143 9  20408301.433 9 -20430454.887 9 -14777553.807 9        214.022
       166.771
  24189931.963 5  24189938.551 5  -5718742.244 5  -4321438.600 5        303.064
       236.145
  23466415.781 5  23466422.123 6  -7157669.216 6  -5420564.840 6      -1972.672
     -1537.149
  21990865.100 9  21990870.047 9  -5776160.908 9  -4352051.799 9      -3197.496
     -2491.555
  24914892.415 5                    -51341.915 1                       3019.125

  25477673.834 4                    -54512.577 1                       3206.270
```

Table 2: Detected Difference for Station ZIMM